1.

(1) 由 $f(y) \le f(x) + \nabla f(x)^T (y-x) + \frac{L}{2}\|y-x\|_2^2$, for all x, y

有 $f(x_{k+1}) \le f(x_k) + \nabla f(x_k)^T (x_{k+1} - x_k) + \frac{L}{2}\|x_{k+1} - x_k\|_2^2$

即 $f(x_{k+1}) \le f(x_k) - t\nabla f(x_k)^T \nabla f(x_k) + \frac{L}{2}\|t\nabla f(x_k)\|_2^2$

$f(x_{k+1}) \le f(x_k) - (1-\frac{Lt}{2})t\|\nabla f(x_k)\|_2^2$

得证

(2) 由(1)有 $(1-\frac{Lt}{2})t\|\nabla f(x_k)\|_2^2 \le f(x_k) - f(x_{k+1})$

$\Rightarrow \|\nabla f(x_k)\|_2^2 \le \frac{2}{t(2-Lt)}(f(x_k) - f(x_{k+1}))$

由 $t \le \frac{1}{L}$ 有 $\|\nabla f(x_k)\|_2^2 \le \frac{2}{t}(f(x_k) - f(x_{k+1}))$

得证

(3) $\|\nabla f(x_0)\|_2^2 \le \frac{2}{t}(f(x_0) - f(x_1))$

$\|\nabla f(x_1)\|_2^2 \le \frac{2}{t}(f(x_1) - f(x_2))$

$\vdots$

$\|\nabla f(x_k)\|_2^2 \le \frac{2}{t}(f(x_k) - f(x_{k+1}))$

求和得到 $\sum_{i=0}^{k} \|\nabla f(x_i)\|_2^2 \le \frac{2}{t}(f(x_0) - f(x_{k+1}))$

而 $\lim_{k\to\infty} x_k = x^*$, 故 $\sum_{i=0}^{k} \|\nabla f(x_i)\|_2^2 \le \frac{2}{t}(f(x_0) - f(x^*))$

得证

(4) $\min_{i=0,\cdots,k} \|\nabla f(x_i)\|_2 \le \sqrt{\frac{1}{k+1}\sum_{i=0}^{k}\|\nabla f(x_i)\|_2^2} \le \sqrt{\frac{2}{t(k+1)}(f(x_0) - f(x^*))}$

得证

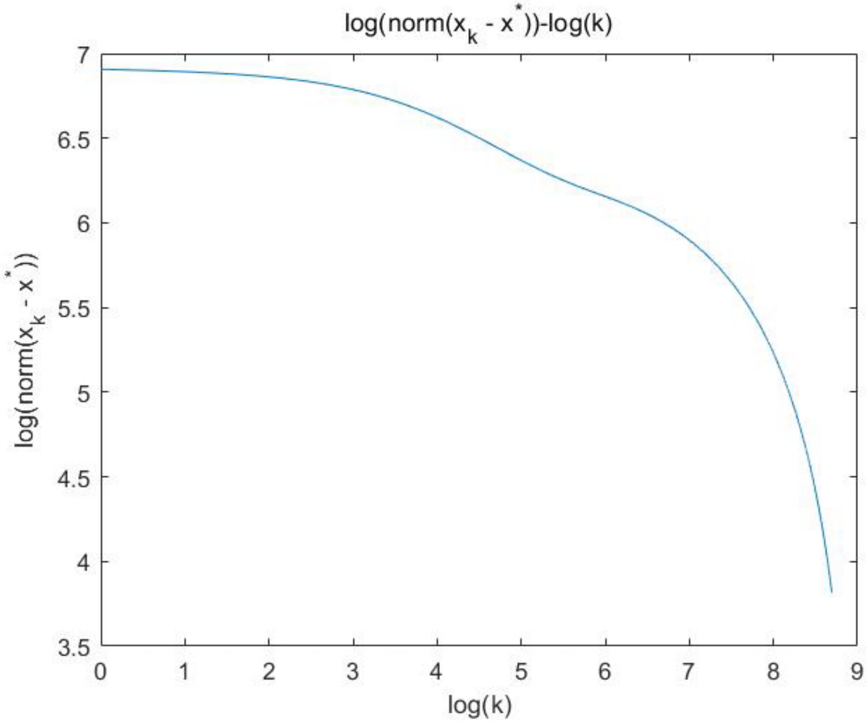
2、

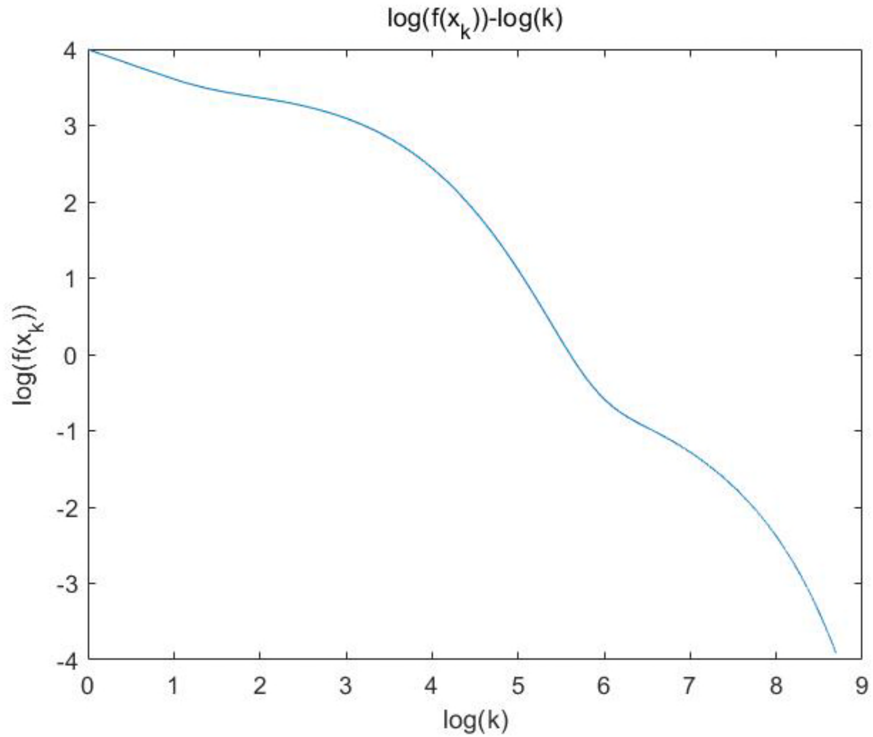
(1) $\mathrm{Prox}_{\alpha f}(v) = \arg\min_x \frac{\alpha}{2}\|Ax-b\|^2 + \frac{1}{2}\|x-v\|^2$

$\Rightarrow \alpha A^T(Ax-b) + x - v = 0 \Rightarrow x = (A^TA + \frac{1}{\alpha}I)^{-1}(A^Tb + \frac{v}{\alpha})$

$\Rightarrow x_{k+1} = (A^TA + \frac{1}{\alpha}I)^{-1}(A^Tb + \frac{x_k}{\alpha}) = x_k + (A^TA + \frac{1}{\alpha}I)^{-1}A^T(b - Ax_k)$

(2) 选取$\alpha = 100$，最终迭代次数$k=5966$，$f(x_k) = 0.0200$，取$f(x) < 10^{-2}$处的x为x^*，图像如下：



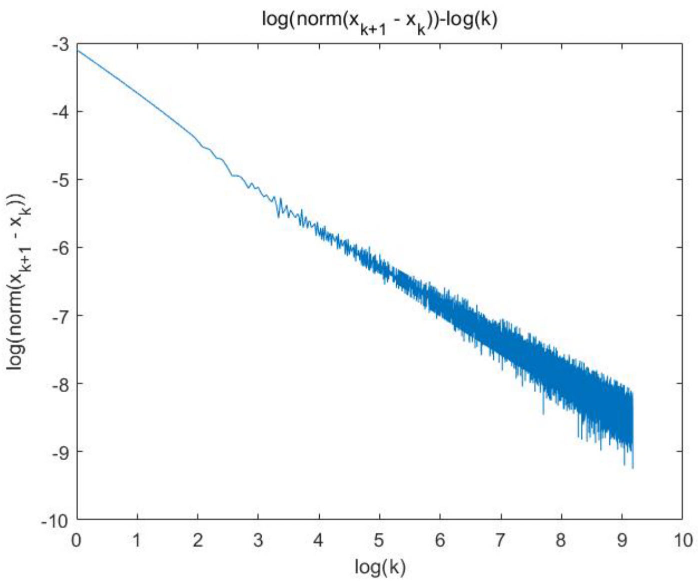


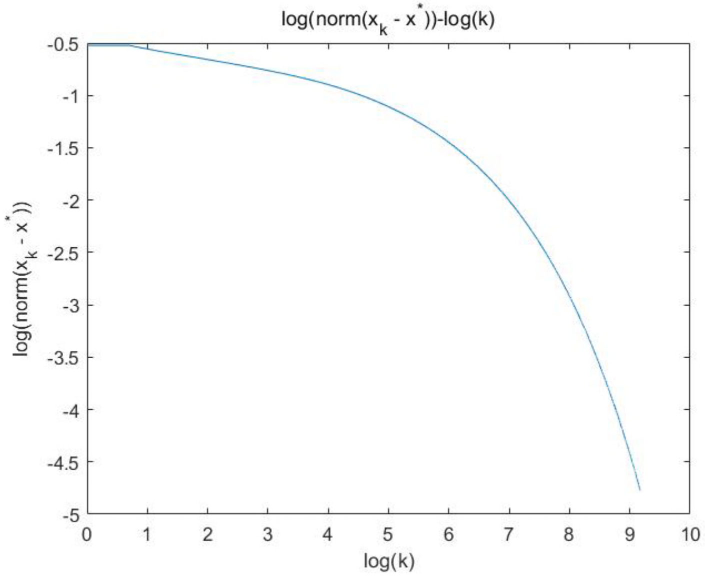
3.

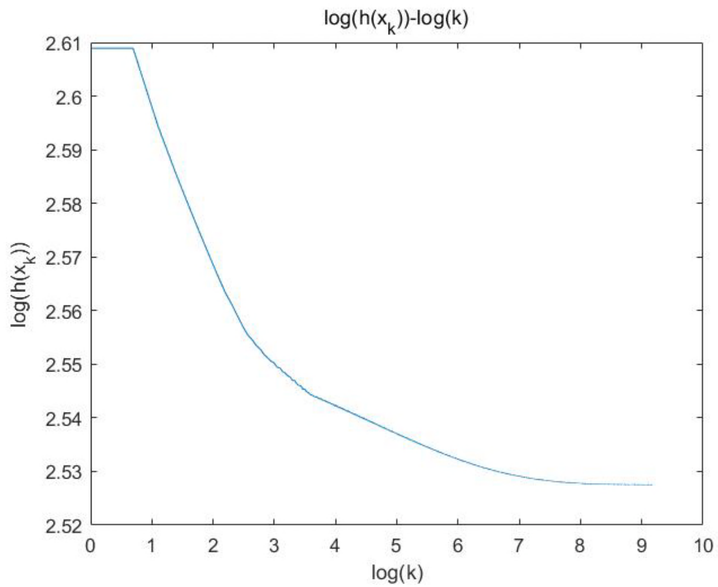
(1) $\partial h(x) = A^T(Ax-b) + \partial\|x\|_1 = A^T(Ax-b) + \begin{cases} 1 & ,x_i>0 \\ -1 & ,x_i<0 \\ [-1,1] & ,x_i=0 \end{cases}$

令 $g_k = \begin{cases} \text{sign}(x_k) & ,\ \text{abs}(x_k) \geq 1e-5 \\ 1-2\text{rand}(\text{length}(\text{find}(x==0)),1), & \text{abs}(x_k) < 1e-5 \end{cases}$

最终迭代次数 $k=9632$，$h(x_k)=12.5225$，选取 $\|x_{k+1}-x_k\|^2<1e-9$ 处的 x_{k+1} 为 x^*，图像如下:

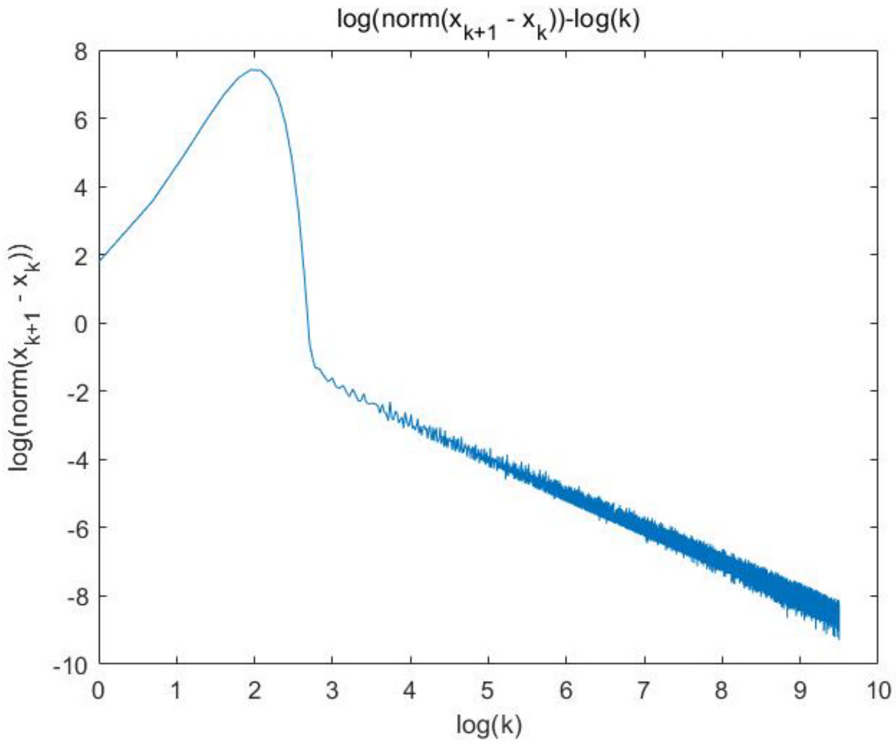


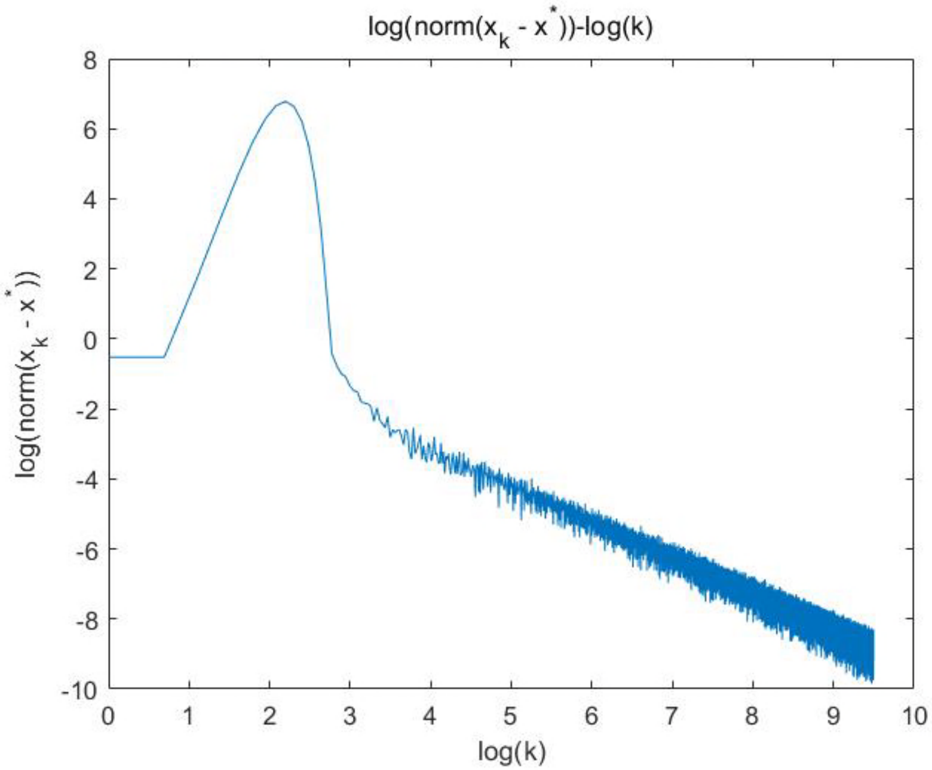


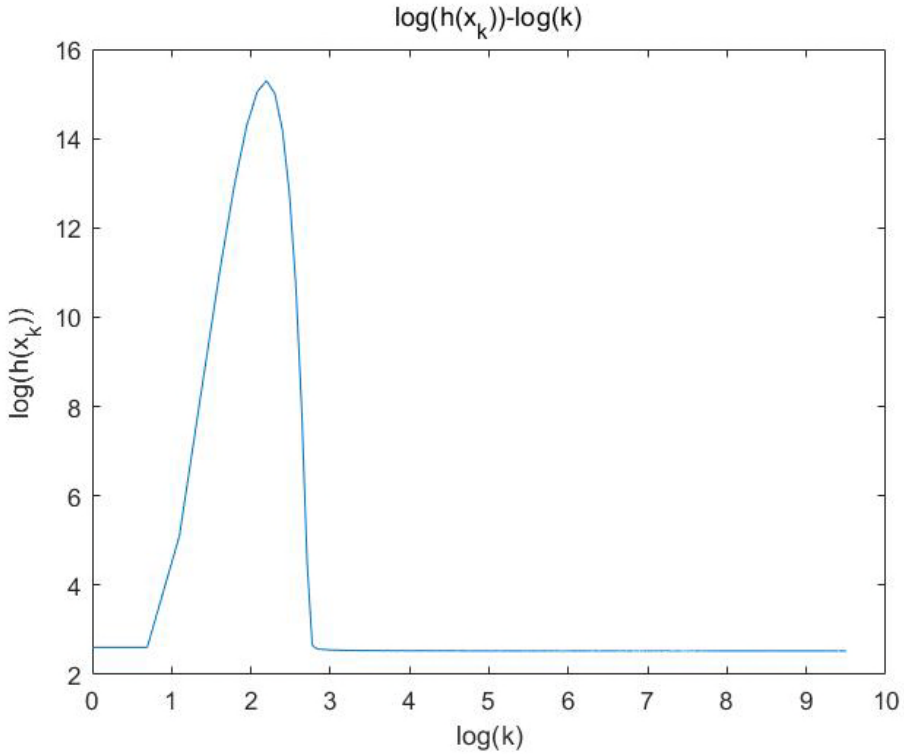


(2) m 为 A^TA 的最小特征值，$m=0.7569$

最终迭代次数 $k=13335$，$h(x_k)=12.5225$，选取 $\|x_{k+1}-x_k\|^2<1e-9$ 处的 x_{k+1} 为 x^*，图像如下:


log(norm(x_{k+1} - x_k))-log(k)
log(norm(x_{k+1} - x_k))
log(k)
8
6
4
2
0
-2
-4
-6
-8
-10
0
1
2
3
4
5
6
7
8
9
10


log(norm(x_k - x^*))-log(k)
log(norm(x_k - x^*))
log(k)
8
6
4
2
0
-2
-4
-6
-8
-10
0
1
2
3
4
5
6
7
8
9
10


log(h(x_k))-log(k)
log(h(x_k))
log(k)
16
14
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
10